**Project Prospectus**

Title: Using suitability analysis to determine the best location for a new Costco location in the Twin Cities Metropolitan Area, MN

Notice: Dr. Bryan Runck

Author: Maisong

Date: 2/17/2021

**Project Repository:**<https://github.com/msongfrancis/gis5572project.git>

**Abstract**

GIS has been used to solve location problems such as finding the best location for retail, emergency services or warehouses (Church 2002). In retail, demographics of existing and ideal costumers are also considered when determining the next retail location (Trubint et al. 2006). Costco, a warehouse style store, has nine locations in the seven metropolitan counties of Minnesota. This study will aim to determine suitable locations for a new Costco location based on social demographics, accessibility, and proximity to competing business and existing Costco locations. The results for potential locations are expected to be similar to existing locations.

**Problem Statement**

*Describe the specific problem and the context. Provide an illustrative figure and/or context map here. In the table, translate the qualitative problem statement elements into specific requirements for the analysis.*

When determining where to build their next retail locations, businesses consider many factors. A few are accessibility, existing services, and demographics of their ideal and existing consumers (Trubint et al. 2006). The Costco i This study aims to determine potential locations for a new retail Costco Store using suitability analysis.

*Table 1. Components needed to find suitable Costco locations.*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **#** | **Requirement** | **Defined As** | **Spatial Data** | **Attribute Data** | **Dataset** | **Preparation** |
| 1 | Road network | Accessible roads – will be used to access proximity to major roads/highways | Road geometry |  | [Mn Geospatial Commons](https://gisdata.mn.gov/dataset/us-mn-state-metc-trans-fnctnl-cls-rds) |  |
| 2 | Areas available for business development | Regionally planned areas for different categories of land use like business development. | Raster | Planning Code | [Mn Geospatial Commons](https://gisdata.mn.gov/dataset/us-mn-state-metc-plan-pland-land-use) |  |
| 3 | Competing Businesses and existing Costco stores | Locations of other businesses that offer similar services like Target, Walmart, and Sam’s Club and existing Costco stores | Point | Lat, long  Business name | Google places API | Transform x,y to points |
| 4 | Social demographics | Social information about each census tract |  | Age, number in household, in work force, median household income | [US Census Bureau](https://www.census.gov/data.html) – ACS survey | Remove unnecessary columns, normalize. |

**Input Data**

The demographic data required are metrics regarding median household income, household size, labor status in the work force and age. The demographic data will be obtained on a census tract level and the census tract boundaries will be sourced from the U.S. Census Bureau. As for the competing locations and existing locations, the coordinates will be obtained from Google using their Google places API. The roads and business development zones will be sourced from MN Geospatial commons as shapefiles.

*Table 2. <insert caption>*

|  |  |  |  |
| --- | --- | --- | --- |
| **#** | **Title** | **Purpose in Analysis** | **Link to Source** |
| 1 | MN Major Roads and Highways | Accessible roads – will be used to access proximity to major roads/highways | [Mn Geospatial Commons](https://gisdata.mn.gov/dataset/us-mn-state-metc-trans-fnctnl-cls-rds) |
| 2 | Regionally planned areas | Regionally planned areas for different categories of land use like business development. | [Mn Geospatial Commons](https://gisdata.mn.gov/dataset/us-mn-state-metc-plan-pland-land-use) |
| 3 | Competing Business Locations | Locations of other businesses that offer similar services like Target, Walmart, and Sam’s Club | Google places API |
| 4 | Social demographics | Social information about each census tract | [US Census Bureau](https://www.census.gov/data.html) |
| 5 | Census Tract | Boundaries for mapping and used for site selection. May calculate centroid for network analysis for accessibility. 2018 | [US Census Bureau](https://www.census.gov/geographies/mapping-files/time-series/geo/carto-boundary-file.html) |
|  | Existing Costco Locations | Examine demographics around existing Costco and how far they are from each other. | Google places API |

**Methods**

*Demographics*

First unnecessary data will be removed form the tables or combine to get total counts. For age, the median age will be used and any tracts with a median age greater than or equal to 26 will be given a one. For median household size, any tract with greater than or equal to 2 occupants will be given a one. For workforce, the total of those in the work force as employed will be totaled and the employment percentage will be calculated for each tract. Then using ranges, a suitability score will be assigned with higher employment percentage getting a value closer to one. Similarly, with median housing income, income within the middle-class range will be given a one and less or above is zero. The median housing income range will come from Pew Research (Cortland 2020). Once every demographic metric has been scored for each census tracts, the tables will be joined to the census tract boundaries and the total score for each tract will be calculated.

\*Not sure about this part\* The census tracts will be rasterized to create raster where each cell has a demographic suitability score. This will be used later to do raster calculations.

*MN Major Roads and Highways and Existing Competing business locations*

Both datasets will be treated similarly. A raster surface will be created showing the Euclidean distance to these features. Then the values will be reclassified. For roads, cells closer to the roads will get a higher suitability value, and for existing competing businesses cells closer to the business will get a lower suitability score ranging from 0-1.

*Regionally planned areas*

These are planned regional development in the metropolitan counties. There are classifications for business and unplanned. I’m not 100% sure if I will use this dataset, but if it were used areas with planned business development would be 1s and the other classes will be given a zero.

Once all rasters for demographic and proximity to major roads and competing businesses are created, I will use raster calculator to add the cells together and output a final suitability raster. Cells with higher values will be more suitable for the next Costco location.

*Existing Costco locations*

With the final suitability raster, I will see if the existing Costco locations are considered most suitable.

*Figure 1. Potential Workflow diagram for analysis.*

*Diagram

Description automatically generated*

**Results**

I expect to get a raster with fuzzy suitability where I can distinguish high, medium, low suitability (Fig 2) and a feature class showing the service area of existing Costco stores - perhaps a 5 miles service area (Fig 3).

Diagram

Description automatically generatedFigure 2. Rough drawing of fuzzy suitability expected from overlaying all rasters.

Figure 3. Rough drawing of expected thiessen polygons to visualize service areas of existing Costco locations.

Diagram

Description automatically generated

**Results Verification**

* Examine demographics around existing locations. Are they similar? Where do they differ?
  + Would expect the existing locations to be similar in demographics.
* Aerial imagery verification. I think a good way to ensure I’ve found a place not a neighborhood or something like a lake feature is to reference the location on imagery.

**Discussion and Conclusion**

* The results should return places best for a new Costco development based on demographics and proximity to roads and other businesses and existing Costco locations.
* Would the new potential locations service the same areas as existing locations? One way to also examine this is to perform a spatial interaction analysis based on the resource availability of the existing Costco locations.

**References**

Church, R.L. 2002. Geographic information systems and location science. *Computers and Operations Research* 29:541-562.

Costco Wholesale Corporation. (2017). Seeking Alpha. https://seekingalpha.com/symbol/COST/earnings/transcripts

Cortland, P. 2020. Are you part of the middle class? This calculator can tell you. Considerable. https://www.considerable.com/money/economy/what-is-middle-class-in-every-state/

Trubint, N., Ostojic, L., and Bojovic, N. 2006. Determining an optimal retail location by using GIS. *Yugoslav Journal of Operations Research*. 16.2:253-264.

**Self-score**

*Fill out this rubric for yourself and include it in your lab report. The same rubric will be used to generate a grade in proportion to the points assigned in the syllabus to the assignment.*

|  |  |  |  |
| --- | --- | --- | --- |
| **Category** | **Description** | **Points Possible** | **Score** |
| **Structural Elements** | All elements of a lab report are included **(2 points each)**:  Title, Notice: Dr. Bryan Runck, Author, Project Repository, Date, Abstract, Problem Statement, Input Data w/ tables, Methods w/ Data, Flow Diagrams, Results, Results Verification, Discussion and Conclusion, References in common format, Self-score | 28 |  |
| **Clarity of Content** | Each element above is executed at a professional level so that someone can understand the goal, data, methods, results, and their validity and implications in a 5 minute reading at a cursory-level, and in a 30 minute meeting at a deep level **(12 points)**. There is a clear connection from data to results to discussion and conclusion **(12 points)**. | 24 |  |
| **Reproducibility** | Results are completely reproducible by someone with basic GIS training. There is no ambiguity in data flow or rationale for data operations. Every step is documented and justified. | 28 |  |
| **Verification** | Results are correct in that they have been verified in comparison to some standard. The standard is clearly stated **(10 points)**, the method of comparison is clearly stated **(5 points)**, and the result of verification is clearly stated **(5 points)**. | 20 |  |
|  |  | 100 |  |